



A physical map of chromosome 2 of *Arabidopsis thaliana*.

E A Zachgo, M L Wang, J Dewdney, et al.

Genome Res. 1996 6: 19-25

Access the most recent version at doi:[10.1101/gr.6.1.19](https://doi.org/10.1101/gr.6.1.19)

References This article cites 25 articles, 7 of which can be accessed free at:
<http://genome.cshlp.org/content/6/1/19.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Copyright © Cold Spring Harbor Laboratory Press

RESEARCH

A Physical Map of Chromosome 2 of *Arabidopsis thaliana*

Eve Ann Zachgo,^{2,4} Ming Li Wang,^{1,2,4} Julia Dewdney,^{1,2}
David Bouchez,³ Christine Camilleri,³ Stephen Belmonte,² Lu Huang,²
Maureen Dolan,² and Howard M. Goodman^{1,2,5}

¹Department of Genetics, Harvard Medical School and ²Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114; ³Laboratoire de Biologie Cellulaire, Institut National de la Recherche Agronomique (INRA), 78026 Versailles CEDEX, France

A yeast artificial chromosome (YAC) physical map of chromosome 2 of *Arabidopsis thaliana* has been constructed by hybridization of 69 DNA markers and 61 YAC end probes to gridded arrays of YAC clones. Thirty-four YACs in four contigs define the chromosome. Complete closure of the map was not attained because some regions of the chromosome were repetitive or were not represented in the YAC library. Based on the sizes of the YACs and their coverage of the chromosome, the length of chromosome 2 is estimated to be at least 18 Mb. These data provide the means for immediately identifying the YACs containing a genetic locus mapped on *Arabidopsis* chromosome 2.

The small flowering plant *Arabidopsis thaliana* is an excellent model system for metabolic, genetic, and developmental studies in plants. Its haploid nuclear genome is small (~100 Mb), consisting of five chromosomes with a low frequency of repetitive DNA (Meyerowitz 1994). A classic genetic map (Koornneef et al. 1983) and maps based on restriction fragment length polymorphisms (RFLPs) (Chang et al. 1988; Nam et al. 1989; Hauge et al. 1993; Lister and Dean 1993) have been published. A cosmid-based physical map containing ~750 contigs obtained by DNA fingerprinting (Hauge and Goodman 1992, AAtdB@<http://probe.nalusda.gov:8300/>) is available, and ~30% of the genome was covered by anchoring the then available yeast artificial chromosomes (YACs) to DNA-based genetic markers (Ward and Jen 1990; Grill and Somerville 1991; Hwang et al. 1991).

A more complete physical map of the *Arabidopsis* genome would be highly advantageous for cloning genetic loci of interest, studying chromosome structure, and sequencing the genome. Of the five *A. thaliana* chromosomes, chromosome 2 is one of the smallest (16%–17% of the genome) and contains one of the two genomic rDNA clus-

ters (Maluszynska and Heslop-Harrison 1991; Albin 1994; Copenhaver et al. 1995). We present here a YAC contig physical map of chromosome 2 of *A. thaliana*. It is aligned to the genetic map (Lister and Dean 1993) and consists of four sets of overlapping YAC clones covering 90% of the chromosome.

RESULTS AND DISCUSSION

The physical map of chromosome 2 of *A. thaliana* shown in Figure 1 was obtained by the hybridization of 130 DNA probes to the CIC YAC library (Creusot et al. 1995). The entire CIC library (1152 clones) was grown in duplicate in a densely packed grid pattern on a single nylon filter the size of a microtiter plate. Dot colony hybridization of radioactively labeled probes made from chromosome 2 markers to the grids allowed the initial anchoring of YACs to chromosome 2. Many of the probes were RFLP-mapped genomic cosmid clones from the fingerprinted cosmid contig map. Supporting data were obtained by probing grids with additional cosmids from the contig that contained the original RFLP-mapped clone. Other molecular markers were obtained from members of the *Arabidopsis* community.

Positively hybridizing YAC clones identified on the grids were confirmed by hybridization of

⁴These authors contributed equally to this work.

⁵Corresponding author.

E-MAIL goodman@frodo.mgh.harvard.edu; FAX (617) 726-3535.

A. *THALIANA* CHROMOSOME 2 YAC PHYSICAL MAP

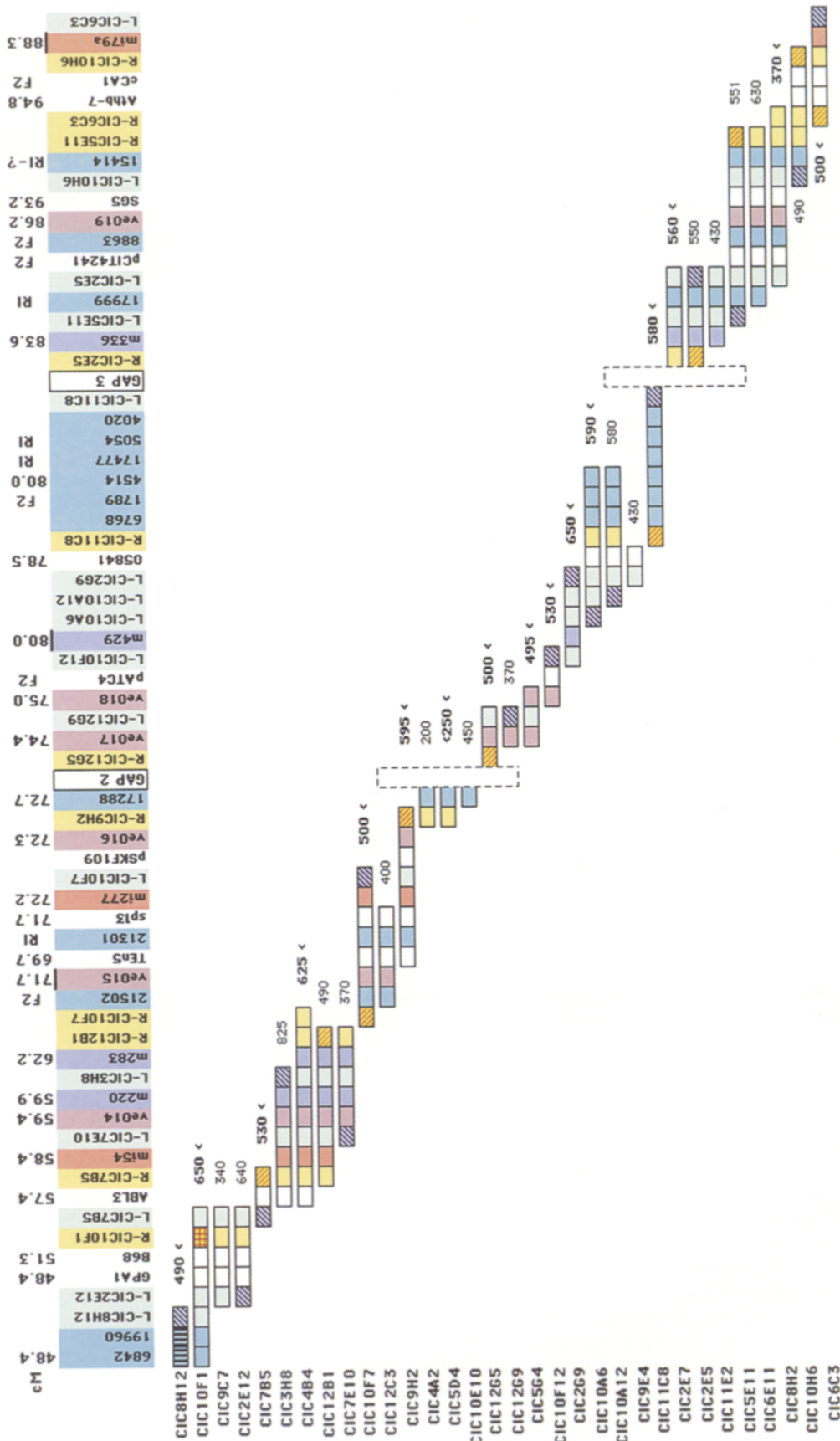


Figure 1 (Continued; see page 22 for legend.)

ZACHGO ET AL.

the probe to Southern blots of restriction enzyme-digested yeast DNA preparations containing the YAC. Comparison of the hybridization patterns of the YACs to *Arabidopsis* genomic DNA verified that the colony hybridization signal resulted from homology of the probe to unique *Arabidopsis* genomic restriction fragments and not cross-hybridizing yeast sequences.

Complete YAC coverage of the chromosome was not achieved by anchoring YACs to the chromosome 2 markers. Therefore, a modification of the inverse polymerase chain reaction (IPCR) was used to generate end probes from the ends of the anchored YACs. Radioactively labeled end probes were hybridized to the grids, and positively hybridizing YAC colonies were confirmed by Southern blot analysis. End probes generated from the YACs enabled one or more steps to be taken to extend the anchored YAC contigs and close gaps between contigs. Fewer than four steps were needed to reach the next anchored YAC or YAC contig. We attribute this to the large insert size

(480-kb mean) and low incidence of chimeric clones in the CIC library (Creusot et al. 1995), as well as to the high number and broad distribution of markers on chromosome 2.

Five regions of the map could not be closed. As depicted in Figure 1, a gap exists to the right of the last YAC clone (CIC6C3) and the telomere at the end of the long arm of the chromosome. On the short arm of the chromosome, the gap between the telomere at the left end and clone CIC7C11 contains a cluster of rDNA genes. CIC7C11 may actually be part of this cluster as it contains rDNA (Creusot et al. 1995). Shown as broken line boxes in Figure 1, three internal gaps exist between mi310 and mi421 (gap 1), 17288 and R-CIC12G5 (gap 2), and L-CIC11C8 and R-CIC2E5 (gap 3). These gaps could not be closed due to the absence of YACs containing these sequences in the library or the repetitive nature of some of the end probes (Fig. 1). The instability of cloned repetitive sequences in yeast makes it difficult to physically map such regions (Schmidt et

Figure 1 A physical map of chromosome 2 of *A. thaliana*. One hundred and thirty DNA probes, hybridized to the CIC YAC library, are ordered along the abscissa. Fifty-eight probes, mapped to chromosome 2 by RFLP analysis, serve as contact points between the genetic and physical maps. The position of markers mapped to the recombinant inbred lines (Lister and Dean 1993) are indicated in cM above the marker names. When a set of probes hybridized to the same YACs, their order was listed by their positions in the genetic map (Lister and Dean 1993). If no genetic or auxiliary data (e.g., cosmid contig; Hauge and Goodman 1992) were available, the exact probe order was indeterminate and was assigned arbitrarily. A line drawn above a marker indicates a difference in marker order between the physical and genetic maps. Markers designated RI have been mapped by us to the recombinant inbred populations and are in correct relative position (with one exception, 15414, marked RI-?) to the markers on either side (no number is indicated as adding markers, and using a different mapping program generates different absolute positions; Hauge et al. 1993); F2 indicates a marker mapped on other mapping populations (Chang et al. 1988; Nam et al. 1989; Hauge et al. 1993) but not the RI lines. Similar types of probes (see Methods) are indicated in the same color: lilac, bacteriophage λ clones (m####); rose, EST sequences (ve012–ve019); peach, genomic clones in pUC119 (mi####); turquoise, cosmid clones (#####); all others are white. Sixty-one probes generated from the ends of anchored YACs are prefixed by L or R to indicate the left (lime) or right (yellow) end of the *Arabidopsis* insert of the YAC. YACs that hybridized to a probe are aligned along the ordinate starting with CIC7C11 (at the top) near the telomere of the short chromosomal arm. The minimal set of 34 YACs are indicated by a < symbol next to the YAC size (bold; in kb indicated adjacent to the YAC). The strings of boxes along the abscissa do not represent the size of the YACs or the exact overlap between YACs but do show the number of probes that hybridized to the YAC. Box color indicates the nature of the hybridization and is the same as the corresponding probe color with the following exceptions: (1) Hybridization of a right end probe to its cognate YAC is indicated by a red sloping pattern on yellow, and a dark-blue sloping pattern on lime for a left end probe, and (2) repetitive sequences—brown, rDNA, green and red, repetitive and/or YAC-specific band patterns. YACs CIC9D11 (Creusot et al. 1995) and CIC6G3 (Creusot et al. 1995; E. Zachgo and M.L. Wang, unpubl.) contain the 180-bp repeat (Martinez-Zapater et al. 1986). Because the ends of two or more YACs may hybridize to the same YAC clone, the diagram may show one of the YAC ends as internal to the rest of the blocks representing the YAC (checkerboard pattern). Broken-line vertical columns indicate the three internal gaps in the map; the double column at Gap 1 represents the uncertainty in the centromeric region. The format of the map requires that a portion of the data for one of the YACs (CIC8H12) and two of the markers (6842 and 19960) be represented in both panels (dark vertical lines on the turquoise background). The entire data set is not depicted here because of space limitations; however, a data supplement may be accessed at <http://weeds.mgh.harvard.edu/goodman/html> on the World-Wide Web.

al. 1994; Creusot et al. 1995). It is likely that gap 1 represents the centromere (Maluszynska and Heslop-Harrison 1991; Copenhaver et al. 1995) because several YACs in this region contain the 180-bp and 5S repeats (Creusot et al. 1995) and other repetitive sequences (Fig. 1; Creusot et al. 1995).

The minimum size of chromosome 2 can be approximated as the sum of the estimated sizes of the four YAC contigs, the gaps, and the telomeres. The data in Figure 1 were used to estimate YAC contig sizes in two ways: (1) Total size (TS) in the four YAC contigs = TYS/AR = 12.8 Mb, where TYS = total YAC size = 36,351 kb and AR = average redundancy = 2.846 (average number of hits per probe); and (2) TS = Σ [YAC size \times (1 - fractional overlap with the previous YAC)] = 13.3 Mb. The fractional overlap with the previous YAC was estimated as the number of markers in common divided by the number of markers associated with the YAC. The sizes of the gaps are unknown. However, using the genetic distance between the markers bordering the internal gaps, and a conversion factor of 200 kb/cM (Koornneef et al. 1983; Meyerowitz 1994), the gap sizes are estimated to be 120–1000 kb, 340 kb, and 720 kb for gaps 1, 2, and 3, respectively. The total size in YAC contigs (~13 Mb), an rDNA cluster of 3.5 Mb (Maluszynska and Heslop-Harrison 1991; Albini 1994; Copenhaver et al. 1995), gaps, and telomeres suggest a minimum size of 18 Mb for chromosome 2.

Virtually complete coverage of chromosome 2 is achieved with a minimal tiling path of 34 YACs (Fig. 1). Contact points between the genetic and physical maps are provided by 58 probes mapped to chromosome 2 by RFLP analysis. The order of the DNA probes in this physical map is generally consistent with the genetic map (Lister and Dean 1993; Fig. 1). Given the low incidence of chimeric clones in the CIC library (Creusot et al. 1995), the map is reliable despite five linkages that exhibit <0.25 units of fractional overlap with an adjacent YAC. Chromosome 2 as described in this study is ~18 Mb in length, 13–15 Mb of which is low-copy DNA and 3.5–4 Mb is rDNA (Maluszynska and Heslop-Harrison 1991; Albini 1994; Copenhaver et al. 1995). Chromosome 2 could therefore encode 2500–3000 genes (5 kb/gene; Meyerowitz 1994).

The YAC physical map of *A. thaliana* chromosome 2 presented here will aid in the isolation and characterization of genetically mapped genes. The map will also provide a framework for

A. THALIANA CHROMOSOME 2 YAC PHYSICAL MAP

the study of plant chromosome structure and large-scale DNA sequencing of a plant genome.

METHODS

Preparation of High-density YAC Grids

Five-by-five grids of the CIC YAC library were generated using the high-density replicating tool of the BIOMEK 1000 Workstation (Beckman Instruments) to spot the yeast clones containing YACs from the 12 plates of the CIC library onto GeneScreen filters. Gridded colonies were outgrown at 30°C (~16 hr) and, after spheroplasting (Matalana et al. 1992), were lysed by autoclaving for 5 min on fast exhaust. Filters were then baked for 2 hr at 80°C and UV cross-linked.

Preparation of DNA

Standard protocols were used to isolate cloned DNA (Ausubel et al. 1987; Sambrook et al. 1989). Total *Arabidopsis* genomic DNA was isolated from ecotype Columbia by standard methods (Watson and Thompson 1986).

Preparation of Probes and Hybridizations

Purified DNA of markers without homology to the pYAC4 vector were random primer labeled, whereas those with homology were either labeled by PCR using universal primers complementary to the vector or the insert was purified and random labeled. DNAs were digested with *Hind*III and labeled with ³²P (Feinberg and Vogelstein 1983) for 1.5 hr at 37°C. End probes were labeled for 16 hr at room temperature followed by 1.5 hr at 37°C after the addition of another 5 units of Klenow. Hybridizations and washes were carried out in glass canisters in a Hybaid oven (National Labnet, Woodbridge, NJ) essentially as described (Hauge and Goodman 1992). Filters were exposed to Reflection film (DuPont NEN) with one Reflection intensifying screen at -80°C for 2–48 hr.

Sources of Probes

The references and/or sources of the various probes used in this study are as follows: cosmid clones (Nam et al. 1989; Hauge and Goodman 1992); bacteriophage λ clones (Chang et al. 1988); EST sequences (D. Bouchez and C. Camilleri, unpubl.); genomic clones in pUC119 (R. Whittier, Mitsui Plant Biotechnology Research Institute, Tsukuba, Japan); AGL3 (Ma et al. 1991); Athb7 (Soderman et al. 1994); B68 (E. Wisman, Max-Planck-Institut für Zuechtungsforschung, Cologne, Germany); GPA1 (Ma et al. 1990); O5841 and O802F (C. Lister, The John Innes Centre, Norwich, UK); pCIT4241 and pCITd100 (ABRC, Ohio State University, Columbus, OH); ABL3 (Long et al. 1993); PR1 (Ulknies et al. 1992); Ten5 (formerly PR21) (Aarts et al. 1995); pTHY-1 (Lazar et al. 1993); pSKF109 (J. Polacco, University of Missouri, Columbia); spl3 (G. Cardon and P. Huijser, Max-Planck-Institut für Zuechtungsforschung, Cologne, Germany); pATC4 (Nairn et al. 1988); AGL20 (formerly SG5) (S. Gold and M. Yanofsky, unpubl.);

ZACHGO ET AL.

cCA1 (E. Tobin and L. Sun; University of California, Los Angeles and Massachusetts General Hospital, Boston, respectively).

IPCR

IPCR to isolate YAC-end DNA was adapted from Ochman et al. (1988). Reactions were performed using a Perkin-Elmer thermocycler. Primers were designed to generate DNA from both ends of the YAC after digestion with *Hind*III or *Clal*, from the left end only using *EcoRV*, and from the right end only after digestion with *Pst*I or *Dra*I. The sequences of these primers are primer 1, 5'-gccaaagttg-gttaaaggcgc-3'; primer 3, 5'-gcccgatctcaagattacg-3'; primer 4, 5'-ggaagaacgaaggaaggagc-3'; primer 5, 5'-aaactcaacgagc-tggagc-3'; primer 6, 5'-catagtactggcgatgctg-3'; primer 7, 5'-tcatacacggtgcctgactg-3'; primer 8, 5'-gactagtagcttagt-catt-3'; and primer 9, 5'-cgtttcttgagctgggac-3'. Primer pairs used to amplify *Hind*III or *Clal* left ends are primers 7 and 1; *EcoRV* left ends, primers 1 and 6; *Hind*III right ends, primers 5 and 3; *Dra*I right ends, primers 3 and 8; *Pst*I right ends, primers 4 and 3; *Clal* right ends, primers 9 and 3.

ACKNOWLEDGMENTS

We acknowledge former (especially Inhwan Hwang and Brian Hauge) and current (John Morris and Deverie K. Bongard-Pierce) members of the laboratory for their varied contributions to this work and Renate Schmidt and Caroline Dean for helpful discussions. We thank our many colleagues in the *Arabidopsis* community for their generous gifts of the various clones used as probes. This work was supported by a grant from GREG 33/94 to D.B. and from Hoechst AG and the National Science Foundation (750-3412) to H.M.G.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Aarts, M.G.M., P. Corzaan, W.J. Stiekema, and A. Pereira. 1995. A two-element enhancer-inhibitor transposon system in *Arabidopsis thaliana*. *Mol. & Gen. Genet.* **247**: 555–564.

Albini, S.M. 1994. A karyotype of the *Arabidopsis thaliana* genome derived from synaptonemal complex analysis at prophase I of meiosis. *Plant J.* **5**: 665–672.

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1987. *Current protocols in molecular biology*. Wiley and Sons, New York, NY.

Chang, C., J.L. Bowman, A.W. De John, E.S. Lander, and E.M. Meyerowitz. 1988. Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci.* **85**: 6856–6860.

Copenhaver, G.P., J.H. Doelling, J.S. Gens, and C.S.

Pikaard. 1995. Use of RFLPs larger than 100kbp to map the position and internal organization of the nucleolus organized region on chromosome 2 in *Arabidopsis thaliana*. *Plant J.* **7**: 273–286.

Creusot, F., E. Fouilloux, M. Dron, J. Lafleuriel, G. Picard, A. Billault, D. Le Paslier, C.D., M.-E. Chaboute, A. Durr, J. Fleck, C. Gigot, C. Camilleri, C. Bellini, M. Caboche, and D. Bouchez. 1995. The CIC library: A large insert YAC library for genome mapping in *Arabidopsis thaliana*. *Plant J.* **8**: 763–770.

Feinberg, A.P. and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6–13.

Grill, E. and C. Somerville. 1991. Construction and characterization of a yeast artificial chromosome library of *Arabidopsis* which is suitable for chromosome walking. *Mol. & Gen. Genet.* **226**: 484–490.

Hauge, B.M. and H.M. Goodman. 1992. In *Methods in Arabidopsis research* (ed. C. Koncz, N.-H. Chua, and J. Schell), pp. 191–223. World Scientific Publishing Co. Pte. Ltd., Singapore.

Hauge, B.M., S.M. Hanley, S. Cartinhour, J.M. Cherry, H.M. Goodman, M. Koornneef, P. Stam, C. Chang, S. Kempin, and L. Medrano. 1993. An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. *Plant J.* **3**: 745–754.

Hwang, I., T. Kohchi, B.M. Hauge, H.M. Goodman, R. Schmidt, G. Cnops, C. Dean, S. Gibson, K. Iba, and B. Lemieux. 1991. Identification and map position of YAC clones comprising one-third of the *Arabidopsis* genome. *Plant J.* **1**: 367–374.

Koornneef, M., J. Van Eden, C.J. Hanhart, P. Stam, F.J. Braaksma, and W.J. Feenstra. 1983. Linkage map of *Arabidopsis thaliana*. *J. Hered.* **74**: 265–272.

Lazar, G., H. Zhang, and H.M. Goodman. 1993. The origin of the bifunctional dihydrofolate reductase thymidylate synthase isogenes of *Arabidopsis thaliana*. *Plant J.* **3**: 657–668.

Lister, C. and C. Dean. 1993. Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. *Plant J.* **4**: 745–750.

Long, D., M. Martin, E. Sundberg, J. Swinburne, P. Puangsomlee, and G. Coupland. 1993. The maize transposable element system *Ac/Ds* as a mutagen in *Arabidopsis*: Identification of an *albino* mutation induced by *Ds* insertion. *Proc. Natl. Acad. Sci.* **90**: 10370–10374.

Ma, H., M.F. Yanofsky, and E.M. Meyerowitz. 1990. Molecular cloning and characterization of GPA1, a G-protein alpha subunit gene from *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci.* **87**: 3821–3825.

———. 1991. *AGL1-AGL6*, an *Arabidopsis* gene family

A. *THALIANA* CHROMOSOME 2 YAC PHYSICAL MAP

- with similarity to floral homeotic and transcription factor genes. *Genes & Dev.* **5**: 484–495.
- Maluszynska, J. and J.S. Heslop-Harrison. 1991. Localization of tandemly repeated sequences in *Arabidopsis thaliana*. *Plant J.* **1**: 159–166.
- Martinez-Zapater, J.M., M.A. Estelle, and C.R. Somerville. 1986. A highly repeated DNA sequence in *Arabidopsis thaliana*. *Mol. & Gen. Genet.* **204**: 417–423.
- Matallana, E., C.J. Bell, P.J. Dunn, M. Lu, and J.R. Ecker. 1992. In *Methods in Arabidopsis Research* (ed. C. Koncz, N.-H. Chua, and J. Schell), pp. 144–169. World Scientific Publishing Co. Pte. Ltd., Singapore.
- Meyerowitz, E.M. 1994. In *Arabidopsis* (ed. E.M. Meyerowitz and C.R. Somerville), pp. 21–36. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Nairn, C.J., L. Winesett, and R.J. Ferl. 1988. Nucleotide sequence of an actin gene from *Arabidopsis thaliana*. *Gene* **65**: 247–258.
- Nam, H.G., J. Giraudat, B. Den Boer, F. Moonan, W.D.B. Loos, B.M. Hauge, and H.M. Goodman. 1989. Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. *Plant Cell* **1**: 699–705.
- Ochman, H., A.S. Gerber, and D.L. Hartl. 1988. Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**: 621–623.
- Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schmidt, R., J. Putterill, J. West, G. Cnops, F. Robson, G. Coupland, and C. Dean. 1994. Analysis of clones carrying repeated DNA sequences in two YAC libraries of *Arabidopsis thaliana* DNA. *Plant J.* **5**: 735–744.
- Soderman, E., J. Mattsson, M. Svenson, C. Borkird, and P. Engstrom. 1994. Expression patterns of novel genes encoding homeodomain leucine-zipper proteins in *Arabidopsis thaliana*. *Plant Mol. Biol. Rep.* **26**: 145–154.
- Ulknes, S., B. Mauch-Mani, M. Moyer, S. Potter, S. Williams, S. Dincher, D. Chandler, A. Slusarenko, E. Ward, and J. Ryals. 1992. Acquired resistance in *Arabidopsis*. *Plant Cell* **4**: 645–656.
- Ward, E.R. and G.C. Jen. 1990. Isolation of single-copy-sequence clones from a yeast artificial chromosome library of randomly-sheared *Arabidopsis thaliana* DNA. *Plant Mol. Biol.* **14**: 561–568.
- Watson, J.C. and W.F. Thompson. 1986. Purification and restriction endonuclease analysis of plant nuclear DNA. *Methods Enzymol.* **118**: 57–75.

Received October 24, 1995; accepted in revised form December 6, 1995.